

EFFICIENT RETRIEVAL OF UNIFORM RESOURCE LOCATORS

DESCRIPTION

BACKGROUND OF THE INVENTION

Field of the Invention

5 The present invention generally relates to interconnected computers and more particularly, the present invention relates to formatted data distributed on interconnected computers.

Background Description

10 Generally what is known as the Internet or World Wide Web (WWW) is a large scale network of globally connected computers. The Internet places an almost limitless amount of information at the fingertips of each connected computer user. The information is located at various globally connected computers, each at what is typically referred to as a "Web Site." Each web site may include multiple accessible web pages linked to each other. Each web page is a data file encoded in HyperText Markup
15 Language (HTML) or a similar browser friendly code. A web page may include one or more Hypertext links or Hyperlinks, each to a Universal or Uniform Resource Locator (URL) that points to a file at an Internet location. The file may be any type of file including but not limited to another HTML document or web page, a word processing document, compressed or uncompressed data or simple text.

Typical URLs specify the location of individual resources at a particular location or domain, but not necessarily in a compact way. An important feature of most domain names are intuitiveness indicating its origin and making memorization easier.

5 Organizations may prefer domain names that reflect or are identical with the organization name, e.g., www.judicialwatch.org. A typical URL is a domain name followed by a path and, then with individual file names attached. So, a URL may become long and unwieldy and under some circumstances may be unusable without some form of compression.

10 These large unwieldy URLs may become troublesome, especially, when large numbers are being handled with limited resources. So, for example, using a wireless web appliance (such as a cellular phone (cell phone) or pager with a built in browser or a personal digital assistant (PDA) connected to the web with a wireless connection) memory, keystrokes and bandwidth may all be at a premium. Clearly, the more larger URLs that are being stored on the wireless web appliance, entered into the appliance's browser and being transmitted between the appliance and its ground station, the more of a problem these large URLs become. This problem is further exacerbated when a URL request from a wireless web appliance results in a display that includes several of these large URLs. Still other circumstances where these large unwieldy URLs may create a nuisance include when the URLs are embedded within some digital contents or in printed bar codes.

20 U.S. Patent No. 5,835,718 to Blewett entitled "URL Rewriting Pseudo Proxy Server" teaches a real-time URL re-mapping method directed to the commercial needs of Internet businesses. Since internet advertisers and services charge for information access, customer interactions HTML based content must be recorded. Blewett teaches a method wherein customers interact with such content through a proxy machine that records the interactions. Thus, the customers' web based requests are directed to a proxy machine

that maps that request to a URL where the desired content is located. However, Blewett does not allow any user to register codes for arbitrary URLs and the method of Blewett hides the true URL to further its goal of tracking requesters.

U.S. Patent No. 5,895,471 to King et al., entitled "Providing a Directory of Frequently Used Hyperlinks on a Remote Server" teaches a directory server that stores directories associated with remote devices. As described in King et al., remote devices include storage, a display and a button and can receive and display hypermedia information based on an identifier, i.e., any state of the art internet browser. The devices also can send identifiers to the directory server, which may add them to the associated directory. A device also may request that the server retrieve previously stored information. Thus, the directory server stores bookmarks for the remote device owner and, merely returns the full URL to any device seeking access to the particular file or document located at the URL.

Further, America On Line (AOL) includes a feature in its service referred to as "keywords." Typically, AOL sells keywords to vendors and then, makes the keywords available for information look up. For example, AOL might sell "travel" as a keyword to a travel agency. So that when an AOL member enters the keyword "travel" in the AOL service, the member is directed to information at the travel agency's site. Normally keywords are short mnemonics that are easy to remember, something much shorter than a URL. This makes keywords easy to distribute, the vendor tells customers the keyword for later retrieval of information. Further, keywords are easily conveyed to an audience in radio or TV broadcasts or in print advertising. However, a human must manually enter keywords.

Accordingly, there is a need for efficient ways of identifying and selecting web based information that do not overload browser resources and seamlessly providing the selected web based information.

SUMMARY OF THE INVENTION

5 It is therefore a purpose of the present invention to reduce the size of web browser cached data;

It is another purpose of the invention to reduce web browser load;

It is yet another purpose of the invention to reduce internet traffic.

10 The present invention is a system, method and computer program product for providing links to remotely located information in a network of remotely connected computers. A uniform resource locator (URL) is registered with a server. A shorthand link is associated with the registered URL. The associated shorthand link and URL are logged in a registry database. When a request is received for a shorthand link, the registry database is searched for an associated URL. If the shorthand link is found to be
15 associated with an URL, the URL is fetched, otherwise an error message is returned.

Alternately, all URLs located at a root page may be listed and associated with shorthand keys or links. Associated files and keys are indexed in an index file. The shorthand codes or keys are combined with the root page to form shorthand URLs.
Requests are made for the shorthand URLs and the index file is searched for the requested
20 shorthand URL. If the requested shorthand URL is encountered, the corresponding file is returned; otherwise, an error message is returned.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, aspects and advantages will be better understood from the following detailed preferred embodiment description with reference to the drawings, in which:

5 Figure 1 is an example of a distributed system for the preferred embodiments of the invention wherein multiply connected users or clients communicate with a database over a network of remotely connected computers;

Figure 2 shows a flow diagram of the steps in registering URL with the server;

10 Figure 3 shows a flow diagram of the steps in retrieving previously registered URLs from the registry database on the server;

Figure 4 is an example of a flow diagram of how shorthand codes are associated in the second preferred embodiment;

Figure 5 shows an example of a flow diagram showing the steps in retrieving a file for a particular URL.

15 DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS OF THE INVENTION

Referring now to the drawings, and more particularly, Figure 1 is an example of a distributed system 100 for the preferred embodiments of the invention wherein multiply connected users or clients 102, 104, 106 communicate with a database 108 over a network 110 of remotely connected computers, such what is known as the Internet or the World Wide Web (www). Figures 2 and 3 represent first embodiment of the present invention wherein an entity, e.g., an on-line shopping service, registers codes with a code server that maintains a shorthand database 108. The registered codes, which correspond Uniform Resource Locators (URLs) for the entity, are mapped in one-to-one fashion to

shorthand codes. When a web browser at one of the remotely connected computers 102, 104, 106 requests access to one of the shorthand codes, the server redirects the web browser to the true URL. As can be seen from the example of Figure 1, the client computers may be, for example, a network computer or thin client 102, a personal computer 104 or, a web enabled cell phone 106 or personal digital assistant (not shown).

Turning to Figures 2 and 3 in this first embodiment, the server functions as a compact code server, maintaining a database 108 of all the URL's that have been associated with a compact code. For example, the original URL, <http://www.sometimes.com/output/novak/novak201.html> (or something even more cumbersome) may be associated with Qa2yZ. Instead of requesting the original URL, a more compact <http://codes.server.com?key=Qa2yZ> may be substituted. The code server, upon receiving a request for the more compact shorthand code would initiate an HTTP redirect to the site corresponding to the original URL. Optionally, the server may charge a nominal fee for each redirect, basing the royalty stream for the service provided upon the number of referrals.

INS. A2 Figure 2 shows a flow diagram 120 of the steps in registering URLs with the server according to the preferred embodiment of the present invention. First, in step 122, the server receives a request from a site owner or, from any web user, to register one or more URLs. Then for each URL, in step 124, the server selects an unused key or shorthand code. Preferably, keys are selected to be as short as is practicable, without being exposed to accidental discovery or to correct, automatically, erroneous entries. With this goal, keys are selected to be sufficiently different from each other such that a number of single character changes are necessary to convert one key to another one. Error correction is effected when an erroneous key is entered, but the error is identifiable, i.e., the key is In step 126, the unused key is entered into the registry database 108 as

X

corresponding to one of the URLs. Also, these short keys may be embedded as hyperlinks in web pages, included in e-mail, and stored in devices with limited storage such as a PDA.

5 Figure 3 shows a flow diagram 130 of the steps in retrieving previously registered URLs from the registry database 108 on the server. First, in step 132, the server receives a shorthand coded request or a key for a corresponding page. Then, in step 134, the server searches the registry database 108 for a URL corresponding to the provided key. In step 136, the server determines whether a corresponding URL is found in the database 108. If the key is not found, e.g., due to an error or from someone entering random sequences of characters, in step 138N, the server returns an error message indicating that the corresponding URL has not been found. Otherwise, in step 138Y, the page corresponding to the key is accessed and forwarded to the requestor.

10

15 Figures 4 and 5 represent a second preferred embodiment of the present invention wherein files at a selected website are reorganized by associating each file with a shorthand code corresponding to its path and location. In this second preferred embodiment, each domain is encoded individually. Instead of a code server as in the first preferred embodiment, codes are associated or assigned individually to the URLs at the participating domain. Code assignment may be at the domain owner's request or, alternately, domain owners may run code assignment software themselves, with database 108 being located at the domain server. A client or web site visitor may suggest or request that the domain owner install the shorthand coding on the domain server to enable or facilitate server access. So, for example, for the domain www.sontimes.com, the file [output/novak/novak20.html](http://www.sontimes.com/output/novak/novak20.html) may be associated, more compactly, with A2gd. The URL <http://www.sontimes.com/A2gd> would be redirected to the appropriate file. Thus, 20 problems associated with long domain names are dissipated transparently to site visitors

25

because, normally, URLs replace IP addresses anyway, e.g., www.sometimes.com may be equivalent to 72.3.109.110.

So, in Figure 4, which is an example of a flow diagram 140 of how shorthand codes are associated in the second preferred embodiment, in the first step 142, a list is created of all files at the root page. Then, in step 144, shorthand keys are assigned to each identified file. In step 146, an index of files with corresponding keys is created.

Figure 5 shows an example of a flow diagram 150 showing the steps in retrieving a file associated with a particular shorthand URL. First, in step 152, a browser requests access to a shorthand URL. Then, in step 154 the domain server searches the local shorthand database 108 to identify a local file to which the requested shorthand URL corresponds. If no corresponding file is identified, then, in step 156N, an error message is returned indicating that there is no corresponding file. Otherwise, in step 156Y, the file identified as corresponding to the requested key is returned.

It should be noted that in the above examples, the part of the shorthand URLs represented by http://codes.server.com part is the same for all shorthand URLs. Accordingly, the amount of information that must be stored or transmitted to refer to the original URL is only the shorthand (Qa2yZ) part. Moreover, the codes.server.com example site optionally may reply with a previously cached page directly or, issue an HTTP redirect as described above. Which option is chosen may depend upon whichever option is more appropriate to the bandwidth or latency requirements of the underlying network at the time of the request.

Further, in contrast to AOL keywords, the present invention uses much shorter free form alpha numeric strings that are designed for transmission by machines and not

5

constrained to an existing English word. Thus, by associating a much larger URL with an arbitrary but much more compact shorthand URL, the present invention realizes optimum compression of storage for keys in the redirection server. In addition, the present invention facilitates servicing competing vendors, since there is no need to favor one vendor over another by assigning one a better (easier to remember) keyword. Moreover, the present invention may be applied to any referral service provider.

While the invention has been described in terms of preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.